

WHAT IS CLAIMED IS:

1. A method for choosing a resource, among a plurality of resources, for servicing a request, comprising:

randomly selecting a first resource among the plurality of resources in accordance with a predefined first random selection function, the first resource having an associated first load value;

comparing the first load value to a predetermined threshold value to determine whether the first load value exceeds the predetermined threshold value; and

upon determining that the first load value does not exceed the predetermined threshold value, assigning the request to the first resource for servicing the request.

2. The method of claim 1, including, upon determining that the first load value exceeds the predetermined threshold value:

randomly selecting a second resource among the plurality of resources in accordance with a predefined second random selection function, the second resource having an associated second load value;

comparing the second load value to the predetermined threshold value to determine whether the second load value exceeds the predetermined threshold value; and

upon determining that the second load value does not exceed the predetermined threshold value, assigning the request to the second resource for servicing the request.

3. The method of claim 2, including, upon determining that the second load value exceeds the predetermined threshold value:

comparing the first load value to the second load value and assigning the request to one of the first resource and second resource having a lower associated load value for servicing the request.

4. The method of claim 1, further comprising:

determining whether the first resource is unavailable for selection;

upon determining that the first resource is unavailable for selection, determining which of the plurality of resources are available for selection and redefining the plurality of

resources to include only those of the plurality of resources that are available for selection, and then randomly selecting a resource from among the redefined plurality of resources.

5. The method of claim 1, wherein the randomly selecting is performed by applying a one-way hashing function to the request to generate a first intermediate value, applying a modulo function to the intermediate value to generate a second intermediate value; and applying a mapping function for mapping the second intermediate value so as to select the first resource from among the plurality of resources.

6. The method of claim 5, wherein the modulo function applied has a modulus whose value corresponds to a total load capacity of the plurality of resources.

7. The method of claim 5, wherein each of the resources of the plurality of resources has an associated respective load capacity; and

the mapping function is a probability density function in which each resource in the plurality of resources has a mapping range whose size corresponds to the respective load capacity associated with the resource.

8. The method of claim 5, wherein each of the resources of the plurality of resources has an associated respective load capacity; and

the mapping function is a weighted mapping function that is weighted in accordance with the respective load capacities associated with the plurality of resources.

9. The method of claim 2, wherein the randomly selecting a second resource is performed by applying a one-way hashing function to the request to generate a first intermediate value, applying a modulo function to the intermediate value to generate a second intermediate value; and applying a mapping function for mapping the second intermediate value so as to select the second resource from among the plurality of resources.

10. The method of claim 9, wherein the modulo function applied has a modulus whose value corresponds to a total load handling capacity of the plurality of resources excluding the first resource.

11. The method of claim 10, wherein each of the resources of the plurality of resources has an associated respective load handling capacity; and

the mapping function is a probability density function in which each resource in the plurality of resources, excluding the first resource, has a mapping range whose size corresponds to the respective load handling capacity associated with the resource.

12. The method of claim 10, wherein each of the resources of the plurality of resources has an associated respective load handling capacity; and

the mapping function is a weighted mapping function that is weighted in accordance with the respective load handling capacities associated with the plurality of resources excluding the first resource.

13. A system for selecting a resource, among a plurality of resources, for servicing a request, comprising:

one or more interfaces for receiving a request and for forwarding the request to a selected resource;

one or more processing units for executing stored procedures;

memory for storing the procedures executed by the one or more processing units, the procedures including:

instructions for randomly selecting a first resource among the plurality of resources in accordance with a predefined first random selection function, the first resource having an associated first load value;

first comparing instructions for comparing the first load value to a predetermined threshold value to determine whether the first load value exceeds the predetermined threshold value; and

instructions for assigning the request to the first resource for servicing the request when execution of the first comparing instructions determines that the first load value does not exceed the predetermined threshold value.

14. The system of claim 13, including second resource selection instructions whose execution by the one or more processing units is invoked when the first comparing

instructions determine that the first load value exceeds the predetermined threshold value, the second resource selection instructions including:

instructions for randomly selecting a second resource among the plurality of resources in accordance with a predefined second random selection function, the second resource having an associated second load value;

second comparing instructions for comparing the second load value to the predetermined threshold value to determine whether the second load value exceeds the predetermined threshold value; and

instructions for assigning the request to the second resource for servicing the request when execution of the second comparing instructions determines that the second load value does not exceed the predetermined threshold value.

15. The system of claim 14, including additional instructions whose execution by the one or more processing units is invoked when the second comparing instructions determine that the second load value exceeds the predetermined threshold value, the additional instructions including:

instructions for comparing the first load value to the second load value and assigning the request to one of the first resource and second resource having a lower associated load value for servicing the request.

16. The system of claim 13, including additional instructions for:

determining whether the first resource is unavailable for selection;  
upon determining that the first resource is unavailable for selection, determining which of the plurality of resources are available for selection and redefining the plurality of resources to include only those of the plurality of resources that are available for selection, and then randomly selecting a resource from among the redefined plurality of resources.

17. The system of claim 13, wherein the instructions for randomly selecting a first resource include instructions for applying a one-way hashing function to the request to generate a first intermediate value, applying a modulo function to the intermediate value to generate a second intermediate value, and applying a mapping function for mapping the

second intermediate value so as to select the first resource from among the plurality of resources.

18. The system of claim 17, wherein the modulo function applied has a modulus whose value corresponds to a total load capacity of the plurality of resources.

19. The system of claim 17, wherein each of the resources of the plurality of resources has an associated respective load capacity; and

the mapping function is a probability density function in which each resource in the plurality of resources has a mapping range whose size corresponds to the respective load capacity associated with the resource.

20. The system of claim 17, wherein each of the resources of the plurality of resources has an associated respective load capacity; and

the mapping function is a weighted mapping function that is weighted in accordance with the respective load capacities associated with the plurality of resources.

21. The system of claim 14, wherein the instructions for randomly selecting a second resource include instructions for applying a one-way hashing function to the request to generate a first intermediate value, applying a modulo function to the intermediate value to generate a second intermediate value, and applying a mapping function for mapping the second intermediate value so as to select the second resource from among the plurality of resources.

22. The system of claim 21, wherein the modulo function applied has a modulus whose value corresponds to a total load handling capacity of the plurality of resources excluding the first resource.

23. The system of claim 22, wherein each of the resources of the plurality of resources has an associated respective load handling capacity; and

the mapping function is a probability density function in which each resource in the plurality of resources, excluding the first resource, has a mapping range whose size corresponds to the respective load handling capacity associated with the resource.

24. The system of claim 22, wherein each of the resources of the plurality of resources has an associated respective load handling capacity; and

the mapping function is a weighted mapping function that is weighted in accordance with the respective load handling capacities associated with the plurality of resources excluding the first resource.

25. A system for selecting a resource, among a plurality of resources, for servicing a request, comprising:

one or more interfaces for receiving a request and for forwarding the request to a selected resource;

control logic configured to randomly select a first resource among the plurality of resources in accordance with a predefined first random selection function, the first resource having an associated first load value, compare the first load value to a predetermined threshold value to determine whether the first load value exceeds the predetermined threshold value, and assign the request to the first resource for servicing the request when the first load value does not exceed the predetermined threshold value.

26. The system of claim 25, wherein the control logic is further configured to randomly select a second resource among the plurality of resources in accordance with a predefined second random selection function when the first load value exceeds the predetermined threshold value, determine whether a second load value associated with the second resource exceeds the predetermined threshold value, and assign the request to the second resource for servicing the request when it is determined that the second load value does not exceed the predetermined threshold value.

27. The system of claim 26, wherein the control logic is further configured to assign the request to whichever of the first and second resources has a lower associated load value when the second load value exceeds the predetermined threshold value.

28. A computer program product for use in conjunction with a computer system having one or more interfaces for receiving a request and for forwarding the request to a selected resource, the computer program product comprising a computer readable storage medium and a computer program mechanism embedded therein, the computer program mechanism comprising:

instructions for randomly selecting a first resource among the plurality of resources in accordance with a predefined first random selection function, the first resource having an associated first load value;

first comparing instructions for comparing the first load value to a predetermined threshold value to determine whether the first load value exceeds the predetermined threshold value; and

instructions for assigning the request to the first resource for servicing the request when execution of the first comparing instructions determines that the first load value does not exceed the predetermined threshold value.

29. The computer program product of claim 28, including second resource selection instructions whose execution by the one or more processing units is invoked when the first comparing instructions determine that the first load value exceeds the predetermined threshold value, the second resource selection instructions including:

instructions for randomly selecting a second resource among the plurality of resources in accordance with a predefined second random selection function, the second resource having an associated second load value;

second comparing instructions for comparing the second load value to the predetermined threshold value to determine whether the second load value exceeds the predetermined threshold value; and

instructions for assigning the request to the second resource for servicing the request when execution of the second comparing instructions determines that the second load value does not exceed the predetermined threshold value.

30. The computer program product of claim 29, including additional instructions whose execution by the one or more processing units is invoked when the second comparing

instructions determine that the second load value exceeds the predetermined threshold value, the additional instructions including:

instructions for comparing the first load value to the second load value and assigning the request to one of the first resource and second resource having a lower associated load value for servicing the request.

31. The computer program product of claim 28, including additional instructions for: determining whether the first resource is unavailable for selection; upon determining that the first resource is unavailable for selection, determining which of the plurality of resources are available for selection and redefining the plurality of resources to include only those of the plurality of resources that are available for selection, and then randomly selecting a resource from among the redefined plurality of resources.

32. The computer program product of claim 28, wherein the instructions for randomly selecting a first resource include instructions for applying a one-way hashing function to the request to generate a first intermediate value, applying a modulo function to the intermediate value to generate a second intermediate value, and applying a mapping function for mapping the second intermediate value so as to select the first resource from among the plurality of resources.

33. The computer program product of claim 32, wherein the modulo function applied has a modulus whose value corresponds to a total load capacity of the plurality of resources.

34. The computer program product of claim 32, wherein each of the resources of the plurality of resources has an associated respective load capacity; and the mapping function is a probability density function in which each resource in the plurality of resources has a mapping range whose size corresponds to the respective load capacity associated with the resource.

35. The computer program product of claim 32, wherein each of the resources of the plurality of resources has an associated respective load capacity; and



the mapping function is a weighted mapping function that is weighted in accordance with the respective load capacities associated with the plurality of resources.

36. The computer program product of claim 29, wherein the instructions for randomly selecting a second resource include instructions for applying a one-way hashing function to the request to generate a first intermediate value, applying a modulo function to the intermediate value to generate a second intermediate value, and applying a mapping function for mapping the second intermediate value so as to select the second resource from among the plurality of resources.

37. The computer program product of claim 36, wherein the modulo function applied has a modulus whose value corresponds to a total load handling capacity of the plurality of resources excluding the first resource.

38. The computer program product of claim 37, wherein each of the resources of the plurality of resources has an associated respective load handling capacity; and the mapping function is a probability density function in which each resource in the plurality of resources, excluding the first resource, has a mapping range whose size corresponds to the respective load handling capacity associated with the resource.

39. The computer program product of claim 37, wherein each of the resources of the plurality of resources has an associated respective load handling capacity; and the mapping function is a weighted mapping function that is weighted in accordance with the respective load handling capacities associated with the plurality of resources excluding the first resource.